Exercise 1 (python code + text):

(a) **Generate** a set $X = \{(y_i, \boldsymbol{x}_i), \boldsymbol{x}_i = [x_{i1}, x_{i2}]^T \in R^2, y_i \in R, i = 1, \dots, 200\}$ from the model

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2 + \eta$$

where η is an i.i.d. normal zero mean noise, with variance 0.05. Use $\theta_0 = 3, \theta_1 = 2, \theta_2 = 1, \theta_3 = 1$. In the sequel, pretend that you do not know the model that generates the data. All you have at your disposal is the data set X.

(b) **Apply** the transformation $\varphi(\boldsymbol{x}) = \begin{bmatrix} \varphi_1(\boldsymbol{x}) \\ \varphi_2(\boldsymbol{x}) \\ \varphi_3(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_1 \cdot x_2 \end{bmatrix}$, on all $\boldsymbol{x}_i$'s of X. Denoting by

$\boldsymbol{x}_i'(\in R^3)$ the image of $\boldsymbol{x}_i$, form a new data set $X' = \{(y_i, \boldsymbol{x}_i'), \boldsymbol{x}_i' \in R^3, y_i \in R, i = 1, \dots, 200\}$.

(c) **Adopting** the linear model assumption in the transformed space and the MSE criterion, estimate the parameters of the model.

Exercise 2 (python code + text):

(a) **Generate** a set $X=\{(y_i,\boldsymbol{x}_i), \boldsymbol{x}_i \in R^2, y_i \in \{-1,+1\}, i=1,\dots,2000\}$, as follows: Select 2000 points in the squared area [-2,2]x[-2,2] of the R^2 space, using the uniform distribution. All points that lie on the positive side of the curve $x_2{}^2 - x_1{}^2 = 0$, are assigned to the class "+1", while all the others are assigned to class "-1". Plot the data using different colors for points from different classes. In the sequel, pretend that you do not know how the data were generated. All you have at your disposal is the data set X.

(b) **Apply** the transformation $\varphi(\boldsymbol{x}) = \begin{bmatrix} \varphi_1(\boldsymbol{x}) \\ \varphi_2(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} x_1^2 \\ x_2^2 \end{bmatrix}$, on all $\boldsymbol{x}_i$'s of X. Denoting by $\boldsymbol{x}_i'$

the image of $\boldsymbol{x}_i$, we form a new data set $X'=\{(y_i,\boldsymbol{x}_i'), i=1,\dots,2000\}$

(c) **Plot** the $\boldsymbol{x}_i'$'s using again different colors for points from different classes and compare the resulting plot with that of (a). Comment on them.

(d) **Adopting** the linear model assumption in the transformed space and the MSE criterion, estimate the parameters of the model.

Exercise 3:

Consider the following nonlinear model:

$$y = 3x_1{}^2 + 4x_2{}^2 + 5x_3{}^2 + 7x_1x_2 + x_1x_3 + 4x_2x_3 - 2x_1 - 3x_2 - 5x_3 + \eta$$

Define a suitable function φ that transforms the problem to a space where the problem of estimating the model becomes linear. What is the dimension of the original and the transformed space?

Exercise 4:

Consider the following two-class nonlinear classification task:

$$x = [x_1, x_2, x_3]^T :$$
$$x_1^2 + 3x_2^2 + 6x_3^2 + x_1x_2 + x_2x_3 > (<)3 \rightarrow x \in \omega_1(\omega_2)$$

Define a suitable function φ that transforms the problem to a space where the problem of estimating the border of the two classes becomes linear. What is the dimension of the original and the transformed space?

Exercise 5:

Consider the following data points $x_1 = [1,1]^T, x_2 = [1,-1]^T, x_3 = [0,0.5]^T, x_4 = [-1,1]^T, x_5 = [-1,-1]^T$, where the first three belong to class $+1$, while the remaining two belong to class -1. Determine the linear classifier that is optimal with respect to the sum of squared errors criterion.

Exercise 6:

Verify the sum, the product and the Bayes rule for the discrete-valued case, using the relative frequency definition of the probability.